RECEIVED

JUL 16 1992

FEDERAL COMMUNICATIONS COMMISSION
OFFICE OF THE SECRETARY

Before the
Federal Communications Commission
Washington, DC 20554

RECEIVED
JUL 16 1992
FCC MAIL BRANCH

In the Matter of

Advanced Television Systems and Their Impact Upon the Existing Television Broadcast Service) MM Docket No. 87-268

ORIGINAL FILE

COMMENTS OF LUCASARTS ENTERTAINMENT COMPANY

LucasArts Entertainment Company and, in particular, its THX Division, submits these comments in response to the Second Report and Order and Further Notice of Proposed Rulemaking ("Futher Notice") in the above-captioned proceeding, FCC 92-174, released May 8, 1992.

1. We believe that the conclusions reached by the ATSC in its deliberations concerning multichannel sound systems for advanced television are correct. In particular, the number of audio loudspeaker channels and their disposition for the most complex member of the hierarchy of channel assignments, namely "5.1" channels, grew out of corresponding work in motion-picture sound. Since a great deal of program material for an advanced television service is expected to be originated as theatrical films, it is important that the ATV system, when used to reproduce the program material which has the highest audio complexity, be capable of faithfully reproducing the complete experience.

2. The 5.1 channel system is well chosen from a psychoacoustic point of view, balanced by the understanding of the channel capacity issues of the medium. The attached paper "Psychoacoustics of Multi-Channel Sound Systems for Advanced Television" gives some details of the background behind the choices made in the 5.1 channel system. It was published in the 1992 NAB HDTV World Conference Proceedings.

3. There are existing films which use a sound track format available only on 70-mm prints called Stereo Surround (i.e., two surround channels in addition to the three screen channels and a low-frequency enhancement channel). These are important films, including *Apocalypse Now, Top Gun, Indiana Jones and the Last Crusade,* and many others for which the proposed 5.1 channel system would be transparent. Thus the 5.1-channel system is supported by existing as well as to-be-produced films.

4. The 5.1 loudspeaker channel system should not be derived by an amplitude-phase matrixing of a smaller number of channels. Amplitude-phase matrixing has had a long history in the film industry under the names Dolby Stereo and Ultra Stereo, but there are quite noticeable artifacts of the process, which have been documented in the attached article "Surround Sound Systems used with Pictures in Cinemas and Homes," which was published in the proceedings of the 1990 Washington Conference of the Audio Engineering Society. Although composite digital coding of multiple channels, using the sound in one loudspeaker channel to mask coding artifacts in other channels may be acceptable, amplitude-phase matrixing is not.

5. The fact that the 5.1 channel system represents only the highest member of a hierarchy of systems, wherein the number of audio channels uses only the bit-rate-capacity that is necessary

No. of Copies rec'd 0+5
List A B C D E

(flexible allocation) is extremely good at handling the problem of resources. Thus a local station could produce its news in monaural sound, and use the part of the data capacity designated for the multiple audio channel system for other services at that time. It could then make a transparent switch to a network film presentation in 5.1 channels, losing only some ancillary data capacity.

Respectfully submitted,

Tomlinson Holman

Tomlinson Holman
Corporate Technical Director
LucasArts Entertainment Company
THX Division
P. O. Box 2009
San Rafael, CA 94912
(415) 662-1900

PSYCHOACOUSTICS OF MULTI-CHANNEL SOUND SYSTEMS FOR ADVANCED TELEVISION

Tomlinson Holman
University of Southern California
School of Cinema-Television
and LucasArts Entertainment Company

ABSTRACT

An advanced television system is expected to have improved sound capabilities, including especially the various effects of stereophony, which are of course related to the number and disposition of the loudspeaker channels employed in the system. This paper examines step-by-step the audible improvements to be expected as the number of discrete channels is increased beyond two, in light of the known psychoacoustics of stereophonic listening. An engineered tradeoff is proposed considering the psychoacoustic improvements gained from additional loudspeaker channels compared to the limitations imposed by the information carrying capacity of the RF channel or alternative media.

INTRODUCTION

The history of stereophony is well known, and the Audio Engineering Society has published an anthology of important work in the field entitled *Stereophonic Techniques,* edited by John Eargle. However, while many people today assume that "stereophonic" means two channels, much of the original work done on stereophony employed more than two channels. Two separate developments of stereophonic sound occurred, one in England and one in the U.S. While Alan D. Blumlein experimented with intensity stereo in England, with its use of coincident directional microphones, a group at Bell Labs under W. B. Snow used spaced omni-directional microphones and corresponding loudspeakers to produce the stereophonic effect. Interestingly, although not widely known, Blumlein had plans and may have made experiments using more channels: he wished to extend stereo from using only the horizontal to using both horizontal and vertical dimensions for sound accompanying motion pictures, involving four coincident microphones and corresponding loudspeaker channels. Thus, at least for centrally located listeners, Blumlein's system could produce sound images anywhere in a continuum from left to right and from up to down, corresponding to the space of the picture (especially in the 1.33:1 aspect ratio picture in use at the time). The Bell Labs work, demonstrated in 1934, used more than two channels, namely three left, center, and right, although they recognized that an infinite number of spaced channels was best. Snow called the use of three channels to represent an infinite number "practical." [1]

The first commercial adoption of stereophonic sound beyond a demonstration occurred during the early 1950's in motion-picture theaters, as the newly introduced invention television threatened to keep people out of theaters. Interestingly, the anamorphic process (Cinemascope™) for wide-screen picture presentation was also invented earlier in the century and not applied until competition made its introduction commercially necessary. The film industry, in seeking a stereophonic system, made use of the Bell research and employed more than two front loudspeakers from the beginning.

Of course, as is well known, two-channel stereophony for home listening got its start with the stereophonic phonograph record in the late 1950's, and grew quickly to include home tape formats, FM broadcasting, the digital compact disc, and television broadcasting. Two-channel stereo is extremely popular because at least some of the beneficial effects of stereophony are easy to perceive even under sub-optimal conditions of both the stereophonic listening condition and the quality of the channel. That is, stereo is easily perceived even from two unassuming loudspeakers while listening off center in a noisy environment (such as listening in a car).

USEFUL EFFECTS OF STEREOPHONY

All stereophonic sound systems are distinguished from monophonic ones by reproducing two effects which are beyond the capability of a single channel and loudspeaker. The first effect is stereophonic imaging, that is, the ability to localize different sources, such as instruments, vocalists, or sound effects, at different positions in space. As the number of sound sources presented simultaneously increases, the positioning of the different elements of sound differently helps to maintain separation of the various sounds. This improves intelligibility in the case of speech, or simply recognition of more sounds simultaneously. The ability to separate the component parts of a stereophonic sound field better than of a monaural one has been known since at least World War II when human factors studies of the performance of people in command and control centers presented with many sounds simultaneously was studied. Note that this effect can be produced by simply having a large number of loudspeakers each reproducing a separate sound, not really a "stereophonic" effect at all, but what one might call a "multiple direction monophonic" effect, to coin a mixed expression.

The second effect is the reproduction of spaciousness. More complex than the first effect, this use of stereophony requires that the sound field be reproduced simultaneously from more than one loudspeaker, unlike the localization effect. It further requires a particular kind of relationship between or among the signals applied to the various loudspeakers: they must differ from one another in certain ways. If two identical signals are applied to two loudspeakers and certain listening conditions are met, then what is perceived is a "phantom image" source lying mid-way between the two loudspeakers. If, on the other hand, the loudspeaker signals differ substantially, but originate from the same source, what may be perceived is a sound field effect having no particular direction, leading to a sensation of spaciousness. Examples of the use of this effect include reproducing the reverberant field portion of a concert-hall music performance, and low-level spatial ambience such as the sound of the corn field rustling in *Field of Dreams*. This effect is the one that is readily perceived even listening off axis, if one were to switch from monophonic to stereophonic reproduction (even with the monophonic sound reproduced over multiple loudspeakers).

So stereophony adds two important effects which may be called imaging and spaciousness to sound reproduction. There is a question however, of how many listeners perceive the one effect vs. the other. For example, a great many people do not take care to seat themselves midway between the loudspeakers in two-channel stereo listening, preferring the convenience of sitting where the furniture dictates. To perceive good imaging from a two-channel system generally requires sitting on the median plane between the two loudspeakers, so a great many listeners are not hearing optimal stereo imaging. Nevertheless they may still perceive the spaciousness of the stereophonic sound field. Thus, of the two, and speaking of current two-channel systems, the spaciousness component of stereophony would have to be given greater weight than the imaging component. Many professional audio personnel will dispute this since it does not correspond to their own experience. They take care to sit in what they call the "stereo sweet spot" while preparing program material, and then hope that the listener at home does the same thing. Perhaps a contrary example would help here. At the twice annual Consumer Electronics Show a great many people come to hear new loudspeaker introductions from manufacturers. Informal observation at these shows demonstrates that even among those supposedly knowledgeable about stereophonic listening, and even in the demonstrations not heavily attended, that most listeners do not bother to seat themselves on the centerline. Even among trained listeners there are those who will bother to make a symmetrical setup, and those who won't, and the latter are certainly in the majority. Thus the enveloping property of stereophony has been given more weight in most actual listening conditions than has the imaging component.

IMPROVEMENTS IN STEREOPHONIC IMAGING

In order to improve stereophonic imaging, several rules may be stated. The first is that reliance on "phantom images," that is, image locations lying between loudspeakers by driving the loudspeakers with identical or proportional signals, does not work very well for most listeners. This is because of the strong action of the precedence effect, which states that localizing sound will occur to the first arriving sound direction, unless later arriving sound is higher in level. Thus if one is seated off center in a two-channel stereophonic listening condition, the closer loudspeaker will predominate, and become the source of localization cues. The best way to get all listeners to perceive the direction as being from the same location is to put a loudspeaker at that location and drive it alone.

Second, we cannot forget that the sound system we are trying to establish is one for producing sound accompanying pictures. Since the picture only covers a small part of the total visual environment, a kind of "window to the world," there is a very different view that should be taken of sound that is meant to accompany and synchronize to the picture, and sound which is meant to be reproduced in the rest of the space. I argue that sound imaging is most important in the area of the picture, for it is here that there is a hope for a match between picture and sound, whereas it is actually undesirable to have strong sound imaging off the picture, since that is likely to be only distracting. An example occurs in the original stereophonic mix of *Ben Hur*. At one point as the protagonist returns to his house, a cow is heard mooing loudly from the "effects" channel (the 1950's designation for what we call the surround channel today). As a boy I wondered what a cow was doing loose in the auditorium! Interestingly, when the film was remixed for re-release in recent years, this "problem" was corrected by moving the cow to a screen channel: the state of the aesthetic art in the use of sound imaging is better in 1990 than in the 1950's.

Since localizing sound to match picture well is desirable, let us turn our attention to how sound imaging can be made to match the picture. First, let us take a common case, that of a conventional NTSC direct-view set accompanied by a two-channel stereo sound system, as there are many listeners using this condition today. If they have followed the rule-of-thumb for conventional stereophonic listening, the loudspeakers and the listening position form an equilateral triangle when seen in plan view. The television set occupies the center of the line between the speakers. The listener sits on axis. In this condition, with sound sources panned to the center, and with matched loudspeakers and symmetrical room conditions, the listener perceives center panned dialog from the center of the line between the two loudspeakers, that is, at the picture. Sounds panned more widely rapidly leave the screen and become incapable of synchronizing to action on the screen. This is why most stereo produced for television display puts all dialog and on-screen sound effects in the center of the stereo field, leaving only ambience effects including audience reaction, and music, to violate the screen boundaries by being wider than the picture.

Advanced television helps in this area, because instead of being viewed as sharper television, I believe it should be viewed as bigger television. Any size increase in the area subtended to the viewer's eye by the television picture will cover more of the stereophonic audio sound field and thus there will be a better chance for a match between sound and picture.

A PROBLEM IN TWO-CHANNEL STEREO

There is a problem in two-channel stereophony that affects all two-channel listening including current television stereo. The problem is prominent even for the ideal conditions of listening that we have stated above. There is a fundamental difference in the nature of a phantom sound image, and one that is produced by a single loudspeaker located at the position of the phantom. A centered, perfect phantom image is constructed of two identical sound fields produced by two loudspeakers. As the sound arrives from the left loudspeaker to the left ear before the sound arrives to the left ear from the right loudspeaker, the time delay between the two sound fields will cause a primary null to occur at a frequency related to the head size. The frequency of the null is about 2 kHz. The null is made broader and shallower by the fact that the head is a large object acoustically and while one sound field has more or less direct access to the outer ear, the other must diffract about the head. Further, different transfer functions arise for the two sound fields in the outer ear (pinnae) due to their different directions of arrival. These effects muddy the waters, spreading out across frequency the effects of what otherwise would be a narrow null: they make it more audible as it occupies more bandwidth (a narrow notch is less audible than a broader dip since it changes the level more in critical bands)! Thus as a sound is panned across a stereo sound field, it will suffer a 2 kHz dip and ripples above that frequency as it is panned from left to center, and the dip and ripples will go away as the pan continues to the right channel. (Acoustical consultant George Augsburger first made me aware of this effect which is also described in [3].)

Thus "perfect" stereophonic listening to two-channel stereo has a defect. Is it audible? I did an experiment at Lucasfilm in which I added a solo singer into the middle of a stereophonic sound field by two different means that I could switch between. While the accompanying music was reproduced in conventional two-channel stereo, and even the reverberation of the soloist was added to only the left and right loudspeakers, the soloist direct sound was switched between a phantom image and a center loudspeaker. The experiment was done in an acoustically symmetrical environment and very careful equalization was carried out using a KEMAR dummy mannikin ear canal response to minimize the level and frequency response differences between phantom and center-loudspeaker imaging. Listeners were constrained to be exactly centered and were not allowed to move. What I found was that first, the phantom image was the best I've ever heard, apparently because so many precautions were taken to make things symmetrical and matched, and 2., that the sound from the hard center loudspeaker was "clearer" than that from the phantom. Thus, for me, even under perfect conditions for two-channel listening, there was a very noticeable improvement in adding a third, center loudspeaker for the front stereophonic sound field.

IMAGING FOR OFF-CENTER LISTENING

Now we have said above that a great many listeners do not listen on the centerline of a two-channel system. They suffer a distortion in the stereophonic imaging component of the sound field, with localization distorting toward the closer loudspeaker due to the precedence effect. Researchers at the German broadcasting research organization IRT have studied the effect of "imaging distortion" by using hidden loudspeakers and on- and off-centerline listening as the number of front loudspeaker channels is increased from two to three to four [2]. In the experiment listeners pointed at the image position for sounds panned anywhere between the extreme left and right loudspeakers across a stereophonic sound field reproduced with a variable number of loudspeakers. The finding was that there is a large improvement in the imaging distortion of off-centerline listening as the number of front loudspeaker channels is increased. There was a large improvement in going from two to three front loudspeakers, and another somewhat smaller but significant gain in going from three to four front loudspeakers.

So it may be concluded that as many front loudspeaker channels as practical is the right number to have: five would make another improvement in stereophonic imaging for off-axis listening, but the improvement would be smaller than in going from three to four channels. Referring back to the fact that we are making a sound system to accompany a picture, we have to be cognizant of the requirements which the picture adds. It has been found through the development of practice that stereophonic panned dialog may be quite often distracting. This speaks to the way that people perceive picture cuts vs. sound ones: all of our lives we have been seeing picture cuts in film and television, and we have a well developed grammar we carry around inside our heads for where we are and where the various characters are as the picture changes. On the other hand, we are not nearly so used to corresponding sound edits. A character who appears screen left in a master shot, then centered in close up "should" have their voice panned to match, but it is often found distracting, so dialog is most often left centered unless the line is supposed to be off screen, or is heard in isolation from lines in time: it is the cuts that are distracting.

Due to this fact of aesthetic life, and due to the phantom image frequency response dip discussed above, I believe that it is essential that the center of the stereophonic sound field be represented by an actual loudspeaker, and not be left to phantom images. While four front loudspeaker channels has an appeal because two wide and two narrow systems would permit wider sound images extended beyond the screen, and a narrower stereophonic image that would match the screen, doing so would still leave the important center of the stereophonic sound field to be a phantom image. So the "right" number of front channels has to be an odd number greater than one. Another factor in this debate is the experience of the film industry, a long-time user of multi-channel sound. Through experience, it has generally been found that three front channels are adequate, even though there is much more emphasis in the film industry on getting the sound field right for listeners seated at many different locations from the picture in order to avoid the stereophonic sweet spot, which is anathema in the egalitarian theater environment. Without a doubt, five front channels would be some improvement in stereophonic imaging for off-axis listeners in the theater, since the positions between left and center, and center and right, depend on phantom imaging in the three-channel system. Of course any improvement which consumes resources should be examined for its cost-to-benefit ratio, and in this case that means the digital bit-rate requirement for the increased number of channels compared to the psychoacoustic benefit to be gained from additional channels. It seems perhaps remarkable that our 1992 view of the world aligns perfectly with Snow's 1934 view: that the use of three front channels is "practical!"

SURROUND SOUND: IMPROVEMENTS IN STEREOPHONIC ENVELOPMENT

The "other" component of a stereophonic sound field I have described above as spaciousness. This word was chosen from a field containing other words such as surround, envelopment, and diffuse-field for a particular reason. A two-channel stereophonic sound system can produce a sense of spaciousness, but it generally speaking cannot produce much of a sense of envelopment. The difference is subtle, but important. Spaciousness refers for example to listening to the characteristic reverberation of a concert hall over a two-channel system. In such a system the entire sound field generally lies between the two loudspeakers, so what is heard is a window of sound opening onto another space, where the reverberation of that space is reproduced through the window. Of course, this is a very unnatural situation since our real-world experience is that diffuse reverberation should appear all around us. In the more general case made possible by a multi-channel system, the enveloping component of the sound field would be reproduced all around a listener, thus supplying the listener with the diffuse-field components of a sound field in a way which is genuinely surrounding.

While some work has been done in the area of reproducing sound images and enveloping sound outside the area between the two loudspeakers of a conventional stereo system using just the two loudspeakers, by far the most robust way to ensure that an enveloping sound field exists is to add surround loudspeakers to the front ones. The advantages of having actual surround loudspeakers include greatly increased freedom from sweet spot effects, and an increased ability to get sounds "hard panned" to the surround channel to stay put where they belong.

Most surround-sound systems, both theatrical ones, and ones for the home, have in the past used only a single monaural channel to carry the surround information. In the theater environment, an attempt is made to decorrelate the sound field arriving at listeners due to the surround sound component by employing many loudspeakers in a surround array. The differing times of arrival of the sound from the various loudspeakers to the listener add comb filtering which tends to diffuse the imaging component of the sound field produced by the surround system, as compared to that produced by the single loudspeakers in the screen channels, and thus to produce a degree of diffuse envelopment from only a single channel. In home systems, the most advanced employ electronic decorrelation of the single source channel into two surround channels for application to left and right surround loudspeakers to achieve the goal of producing a diffuse sound field from the surround system.

What is wrong with only one channel of surround supplied by two or more loudspeakers? Very careful setup and precise centerline listening leads to a bizarre effect: the "surround" component winds up yielding a center-of-the-head earphone-like impression of the sound field. I found this to be true in 1983 while listening to well-aligned dubbing stage sound systems, and Plenge [3] has independently found it to be so in experiments he made in Germany. So loudspeakers can produce a center earphone impression, but this is only true for precise centerline setups. Off center, the precedence effect quickly takes over and produces a surround impression in which the direction is dominated by the earlier arriving sound field. Now neither of these conditions is surround sound: one is middle of the head, and the other is off to one side. While the surround array in the theater, and the electronic decorrelator in the home help to reduce the obviousness of this effect, the best solution is multiple surround channels. The related problem of the best position for surround loudspeakers in the home was studied indirectly by Tohyama and Suzuki whose experimental conclusions demonstrate the utility of side rather than rear, and one rather than two, surround loudspeakers. [4]

Once again, Theile has studied this question experimentally. He tried one vs. two vs. four surround channels, all reproduced by four loudspeakers arranged to the sides of the listening area. Special program material had to be prepared to ensure that there were indeed four independent signal sources for four surround channels accompanying the front channel signals. What was found was once again a very large improvement in the enveloping property of the surround sound as the number of independent channels went from one to two, and another smaller gain as the number increased from two to four.

The Digital Sound on Film subcommittee of the Audio Recording and Reproduction Technology Committee of the SMPTE studied this question several years ago, with an eye toward making a recommendation about the number of channels needed in motion-picture production. Working on the basis of making the optimum tradeoff between the number of channels and the bit-rate-requirements of recording sound on the available space on film digitally led to the conclusion that two surround channels was optimum. This system was to be transparent in channel capacity to those film titles, dating back to *Superman I* and including *Apocalypse Now, Top Gun, Hunt for Red October,* and *Indiana Jones and the Last Crusade,* among others which made use of the stereo surround format on 70 mm prints of these films.

By the way, what is left unstudied is the effect of three vs. two surround channels. While four produce greater envelopment than two, do three do nearly the same job? Three surround channels comes to mind because the use of three would permit two side and one rear channels, allowing for directional effects in which a sound originates in, for example, the front center, and pans through the left and right surrounds to the rear, which could be a desirable effect. Whether the addition of an additional channel just to sustain this effect is a good engineering tradeoff is questionable, and the question of improvements in envelopment as one moves from two *through* three to four surround channels has not been studied experimentally.

So the optimum tradeoff, and the one that corresponds to the same kind of thinking in the screen channels in the number of surround channels is probably two. More channels could produce better localization of surround sound, but this is not generally a desirable property anyway since a hard sound image would be unsupported by picture. More could also produce improved envelopment, but at the cost of more complex production requirements and digital bit-rate capacity of the channel, probably at the level of diminishing returns. The placement of the surround channel loudspeakers in optimally to the sides, not the rear, since moving the loudspeaker channels closer together towards the rear increases their cross-correlation which is not desirable.

Of course, not all program material in which front-stage stereophonic is useful needs a surround sound component. For example, a television talk show without audience would benefit from front-stage stereo, since guests may talk over one another and stereophony would lead to better intelligibility of the multiple sources, whereas an enveloping surround sound component of the sound field seems not very useful. On the other hand, there is plenty of material besides movies in which a surround sound component has already proved its utility in television broadcasting. Events such a sports, shows with an audience, classical music shows, etc., can all make good use of a surround channel.

Further experiments which I made on this and related topics are reported in "New Factors in Sound for Cinema and Television," *J. Audio Engr. Soc.* 1991 July/August.

LOW-FREQUENCY ENHANCEMENT

Some program material which is expected to be broadcast on an advanced television system will certainly be motion-picture originated. A study I made of the sound pressure levels on a number of films with calibrated monitoring showed that the largest signal handling capability is required in the 63 Hz octave band for motion pictures, at a considerably lower frequency than the peak found in much musical program material at 125 - 250 Hz. Since program material from both film, and from occasional musical selections having large amounts of low-frequency energy are both possible and desirable for reproduction over a new system, there should be some provision for reproducing such material as the need arises. The problem with leaving enough headroom in the main channels to accommodate these signals is that the main channels would wind up being undermodulated for a large portion of the time. This is because human hearing requires higher sound pressure level at low frequencies to sound equally loud as the mid-range (the Fletcher-Munson curves are obsolete; ISO standards show the current thinking). Thus "leaving room" for these occasional signals would lead to chronic undermodulation. Since human hearing is not particularly sensitive to direction at low frequencies, and also since the bandwidth and corresponding bit rate needed for a separate low frequency channel is low, the use of one optional low-bandwidth channel to carry extra low-frequency information (not intended to be decoded or used by everyone, but by those having more sophisticated systems) is a useful addition to the total system. This channel requires 1/200 of the bandwidth of one of the main front or surround channels.

PERCEPTUAL DISCRETENESS OF THE CHANNELS

A possible solution to the problem of handling multiple loudspeaker channel information over fewer transmission channels is the use of amplitude-phase matrix encoding and decoding. This process is widely used in recording for films and their corresponding video releases. Yet although it remains a popular and practical system, there are drawbacks which the production community recognizes in the perceptual discreteness of sound imaging over such systems. I have written about this in a paper called "Surround Sound Systems for use in Cinemas and Homes" published in the proceedings of the 1990 Washington Audio Engineering Society Conference.

For a new system which is to last for many years, perceptual discreteness of the loudspeaker channels is assumed as a requirement. This does not necessarily mean independent digital coding of the signals for the various channels however. Methods of composite coding a plurality of channels which are intended to be reproduced simultaneously may well permit bit-rate-reduction greater than that offered by two-channel coding schemes. This is because all aggressive bit-rate-reduction schemes rely on perceptual masking as their basis, and the masking need is only to be accomplished thoroughly for listeners equipped with but two ears!

CONCLUSION

Great improvement in stereophonic imaging is made by adding a center channel to the conventional left and right ones, and this improvement is only helped by the presence of an accompanying picture. While more front channels improves off-axis imaging distortion, the number should remain odd, so two more channels are needed to move beyond three and the law of diminishing returns sets in. Thus three front channels seems to be optimal.

Surround channels greatly aid the enveloping property of a stereophonic sound system, and the minimum number that will produce the desired effect without compromise is two. Once again, more surround channels improves envelopment and the capability to individually localize surrounds, but this is not felt to be an extremely important capability. So two surround channels also seems to be optimal.

An added low-frequency enhancement channel used optionally is very useful for some program material which is expected to be broadcast.

Thus the psychoacoustically most useful number of compositely coded loudspeaker channels, traded off against an upper limit set by consideration of the information carrying capacity of a limited bandwidth channel and also by reasonable production cost limitations, is 5.005.

REFERENCES

1. William B. Snow, "Basic Principles of Stereophonic Sound, *J. Soc. of Motion Picture and Television Engrs.*, vol. 61, pp. 567-589.

2. Roland Rebscher and Günther Theile, "Enlarging of the Listening Area by Increasing the Number of Loudspeakers," Audio Engr. Soc., Preprint from the 88th AES Convention, March 1990.

3. G. Plenge, "On the Behaviour of Listeners to Stereophonic Sound Reproduction and the Consequences for the Theory of Sound Perception in a Stereophonic Sound Field," Institut für Rundfunktechnik, München FRG.

4. Mikio Tohyama and Akira Suzuki, "Interaural Cross-Correlation Coefficients in Stereo-Reproduced Sound Fields," *J. Acoust. Soc. Am.*, vol. 85, no. 2, 1989 February, pp. 780-786.

CREDITS

Cinemascope is a trademark of Twentieth Century Fox Film Corporation.

Surround Sound Systems used with Pictures in Cinemas and Homes (Revised)

Tomlinson Holman
University of Southern California
School of Cinema-Television
and
Lucasfilm Ltd.

Abstract

Surround sound technology has grown in recent years to a new level of maturity, with steadily increasing theatrical usage and rapidly increasing home usage. The technology is reviewed in historical perspective with information useful to mixers engaged in using the medium to its greatest effect. One difference due to human perception between the theater and the home is examined.

Background

Reproduced sound for feature films has some interesting similarities and some striking differences from other forms of sound reproduction. Sixty years ago, motion picture sound dominated the whole reproduced sound business, as the sound revolution in the movies caught on so fast that several hundred millions of dollars worth of silent feature films were abandoned (in 1928 dollars!)[1], since patrons no longer wanted to go to movies without the innovation of sound. "Russia's revolution was not more basic," said contemporaries.[2] With the advent of television in the late 1940's, development of film sound surged ahead so that the theater could maintain a technological and sound quality lead over its emerging competition. Stereo sound for film was feasible as early as the '30's, but was not commercially practiced until perceived competition from television forced its introduction with *The Robe* in the early '50's, although there were earlier experiments.

The early period of film sound was a time when the motion picture studios thoroughly dominated the business, owning the means from story development through to exhibition in studio-owned theaters throughout the country. Technically speaking, a strong central studio sound department controlled quality from recording on the set through to exhibition. Although different studios took different approaches to solving a variety of technical problems, and thus made product difficult to interchange between theaters owned by different studios, this was hardly a problem since the prints were never interchanged between competing studio's theaters. Still the overall quality was well controlled within the technological limits of the day. Prescriptions to the projectionist reiterated in 1938[3] reveal that the idea is very old that the projectionist does not hold the last creative job in the sound process, that is, he is to use a constant and correct volume control fader setting.

With the anti-trust forced breakup of the studio's stranglehold over exhibition in the early 1950's, a new era emerged. Since product to be shown at one theater now needed to be compatible from studio to studio, the standardization effort, started in 1917 with work on film perforations by the SMPE, increased in the sound area. Sound-on-film by this time was less proprietary and litigious than it had been, the "invention wars" of the 1930's having largely settled down as the business matured.

Through the later '50's and the '60's, not much innovation occurred, as the number of theater screens decreased in the U.S. through competition from television. Yet beginning in the '70's, the business recovered and a building boom began. Along with this rise, technological innovation found new audiences. Dolby Laboratories improved optical sound tracks by removing the high-frequency filtering called the Academy filter which had been necessary to obscure excessive hiss by replacing it with companding noise reduction, and thus extending the bandwidth by use of a noise reduction system. They also adapted phase matrix technology for a 4:2:4 system since the space on optical sound tracks was limited, and began 1/3-octave-band equalization of hundreds of theaters. The economic impact of a new breed of films such as *Star Wars* and *Close Encounters of the Third Kind* helped greatly to sell the innovation.

The rising importance of standards, when seen with the much earlier ideas of standardizing the fader settings, brought about the primary philosophical difference between film sound reproduction, and that of most of the rest of the audio industry. In film sound, there is no "reality" which is being attempted typically. Instead, the reality is the experience designed to play over a sound system created by the sound editors, mixers, and director of the picture. This differs in philosophy from classical music recording, where the attempt is more or less to transport listeners to the concert hall. When it comes to more popular styles of recording, the film sound method still differs, since standards for reproduction in music studios are much less prominent than in the film business, leading recording studios to compete on how good their monitor systems sound, rather than on how well they achieve particular standards. While not non-existent in the film business, this notion is much less prevalent.

Through a series of explicit standards, combined with many accepted conventions that are not explicitly standardized, the sound field from one studio dubbing stage to another, and from dubbing stage to theater, is noticeably more consistent than in other major areas of recording.

Among the factors which are the subject of explicit or de facto standards are:

1. Electro-acoustic frequency range and response of the sound system measured at listening positions with specified methods* (SMPTE A202M (especially latest revision in committee), ISO 2969),

2. Volume control setting: the playback system fader is calibrated so that a specific indicated mark produces a specific Sound Pressure Level from a known level in the film medium* (for optical sound "50%" modulation of a variable area optical track, and for magnetic sound, 185 nW/m reference fluxivity produces 85 dBC SPL on a slow reading Sound Level Meter for each channel)[4],

3. The arrangement of loudspeakers: three front, within the picture area, left, center, and right; and a set of surround loudspeakers arrayed around the space arranged to produce uniform SPL and a largely horizontal sound field*; and for 70-mm presentation, additional bass handling loudspeakers behind the screen,

4. Screen loudspeakers having directivity control arranged to produce substantially uniform direct fields over the area of audience to be served, but without excess energy directed to other surfaces,

5. Controlled background noise level maximum, reverberation time and its frequency variability, and echoes and other effects of discrete reflections.

Dolby Laboratories played the principal leadership role in the standards marked with a *; Lucasfilm THX in the last one.

Another interesting standardizing force in the film sound business is the annual search for an Academy Award winner for sound. During this search, each of the films under consideration is played under standardized conditions, and the film which sounds best to the auditors wins. This contest is a powerful ally in helping to standardize dubbing stage practice, since if the film sounds much different in the Academy theater than it did in the dubbing stage, questions will certainly be raised about the quality of the dubbing stage sound system and its alignment to the standards.

It should be pointed out that the items standardized above are largely operational in nature, rather than being specifications for quality. Of course, quality counts too, and some of the factors which affect quality do appear in appendixes to standards, and some which are still important, do not. One might say that a necessary but not sufficient condition for good sound quality is the observance of the standards referenced above.

Surround Sound on Motion Pictures

The principal factor which improved the economic chances for wide spread distribution of stereo movies was the introduction of matrixed stereo variable area optical sound on film. In the classic 1950's stereo movies, the manufacture of the 35-mm prints involved two extra steps after exposure and development by the principal laboratory. In the first step, the film was "striped" with four magnetic oxide stripes. This process was difficult and reject prone for a variety of reasons. Second, each reel had to be re-recorded on one of a bank of stripe-track recorders in real time. Since printing of the picture can occur at much faster than real time, but "sounding" of the prints required real-time operation (high speed duplication has not been applied to film since the 5-mil thick acetate base has physical distortions in it which could lead to poor head contact at high speed), the process effectively priced magnetic print manufacture out of the marketplace, except for specialized use on the limited release 70-mm format.

So starting in the 1970's at Dolby Laboratories, methods of encoding and decoding first three, then four channels worth of information into the two available tracks of the optical sound track area were developed. (Using narrower tracks in order to get more of them was deemed too intolerant of projector misalignment.) While the starting point for this technology was certainly the quadraphonic matrix developments that had come in the ten years before, optimization for sound track usage was performed at the outset, and development has continued.

Consider how conventional stereo can be considered an "encoded" system. If we take the arc tangent of the ratio of amplitudes of the left and right channels, we have the angle for a vector of unit length which points in various directions. The vector has its center at the origin of a circle, and it points "west" for left channel only information, "east" for right channel only information, "north" for precisely centered information, and "south" for difference information (L – R). Adding arbitrary phase information extends the possible points from those lying on the unit circle, to those lying on a sphere having the same radius as the circle.

We can then imagine a variety of placements for the four channels on the surface of the sphere. If we consider simply separation of the channels, we would wish to place the four channels as far apart on the sphere as possible. Such a system is called tetrahedral since the object which just encompasses the extremes of such coding has such a shape. All channels are equally far apart from each other. The resulting unenhanced crosstalk is –4.77 dB between all channels. But there are several drawbacks of this system. One is that phase shift networks are necessary in the decoder, which are subject to tolerance considerations for exploitation on a wide-spread basis (phase sensitivity to component variation is generally worse than amplitude sensitivity), and their concomitant size and cost in consumer applications. Also, such an encoding scheme is incompatible with conventional stereo. Thus a hierarchical scheme with conventional two-channel stereo as a member is precluded.

The chosen scheme for the Dolby MP Matrix is to keep the encoding basically the same as it is in two-channel stereo program material: left input goes only to the Left Total track, center input is split between Left Total and Right Total tracks, right input goes only to the Right Total track, and the surround input is split between Left Total and Right Total tracks, with a $\pm 90°$ phase shift relative to Lt and Rt. Although the uncorrected crosstalk of this scheme is –3 dB between adjacent channels, the slight improvement of the tetrahedral system is more than overcome by the compatibility with conventional stereo, making the matrix hierarchical. This means that listened to undecoded, the Left Total and Right Total signals played over left and right loudspeakers are essentially conventional stereo, and played over a mono system, the same advantages and disadvantages of listening to stereo sources in mono accrue, but no worse artifacts occur. These include, of course, cancellation of difference signal (surround) information in mono listening, requiring program producers to think about what part of their program it is possible to give up for the mono listener. This hierarchical compatiblity was important in the theater, since mono compatibility helped sell the format, and is an important consideration in using the scheme for television, where the listener will have unknown conditions, ranging from mono, through stereo, to 4-channel matrix-decoded systems.

The principal design considerations which were used to make the matrix specific to motion pictures were two. The first is that the traditional left/right/front/back four-channel matrix decoded system was rotated to provide three screen channels and one surround channel. The reason for this is the strong emphasis on sound imaging which can match picture imaging on the screen, and the relatively little importance of actually locating sounds that are off the screen. The second is a set of techniques to suppress the leakage of screen sound information into the surround channel, since crosstalk into surround loudspeakers, particularly of the largely center channel dialog, is especially annoying to ordinary viewers, as it "breaks" the edges of the screen, and the concentration of the viewer. These techniques include:

1. Matrix encoder and decoder low-pass filters in the surround channel. Since matrixed technology depends on phase matching of the transmission paths of the LT and RT channels more than for ordinary two-channel stereo, and since phase errors between channels generally grow with increasing frequency, the bandwidth of the surround channel is restricted to about 7 kHz.

2. A matrix encoder high-pass filter. Since surround sound systems in theaters and homes generally use loudspeakers having less low-frequency power handling capacity than screen channels, and since low frequencies contribute relatively little to localization, a 100 Hz high pass filter is employed.

3. A noise-reduction companding technique, based on 1/2 the action of a standard Dolby B type NR compander, is employed. This helps reduce sibilant splatter of dialog "s's" into the surround channel and also helped to reduce the hiss component of the surround channel to be less than that of the screen channels perceptually, thus helping to conceal the presence of the surround loudspeakers. This is necessary due to the differences in perception of the screen channel vs. the surround channel. The desire is to keep the localization of any hiss due to the medium on the screen, thus helping to conceal the surround loudspeakers.

4. A time delay is used in the decoder to make use of the precedence effect. By delaying the arrival of the surround energy behind that of the screen sound by about 20 ms, any crosstalk remaining is reduced perceptually, since the precedence effect will tend to produce sound imaging that coincides with the first arriving energy.

Decoding technology has changed substantially since the first introduction of the professional system. The first theatrical decoders employed gain riding by a few dB responding to the relative levels of voice frequency range energy in the two channels to produce center hardening in a left, center, right only system, and used three variable gain elements. Subsequently, a fourth channel was added by using directional information derived from phase between Lt and Rt to drive a variable matrix, and these decoders used four variable gain elements for the audio.

Contemporary Dolby decoders employ a constant power variable matrix controlled by the relative amplitude and phase of Lt, Rt, Lt + Rt and Lt – Rt. The controlling signal is derived from Lt and Rt signals by bandpass filtering, logarithmic rectification, and derivation of two bipolar signals representing left-right and center-surround dominance. These two signals are detected nearly instantaneously, and then passed through a dual-time-constant RC combination, which provides long time constants for best signal "solidity" (lack of wandering) except when needed by fast changes. The two bipolar signals are then phase split producing a total of 4 control voltages, which are applied to the control ports of a set of 8 voltage controlled amplifiers. Four of the VCA's have Lt as their audio input, and four have Rt. By precision addition and subtraction of the outputs of the 8 VCA's, four outputs are created, L, C, R, and S. The constant power principle using "signal dominance" and crosstalk cancellation by addition of out-of-phase components in the matrix of VCA's and output summers produces high stability of program loudness, since the action of the decoder does not change the signal power. Use of the dominance concept yields subtraction of crosstalk from adjacent channels which is needed since crosstalk in opposing directions is low anyway. Another way to think about how the decoder works is to realize that a line through left, center, and right is produced by relative amplitude differences between Lt and Rt channels, and a second line, perpendicular to the first, is encoded by relative phase differences between the source channels. This means that information encoded within the triangle described by left, right, and surround, but off the two lines described above, is encoded by means of both relative amplitude and relative phase information.

The modern high-quality consumer decoders are licensed under the trade name Pro Logic, and are detailed in Dolby Laboratories Information bulletin "Dolby Pro Logic Surround Decoder Principles of Operation," by Roger Dressler, available from Dolby Laboratories. Although the most modern method uses eight voltage controlled amplifiers, this number has proved possible in practice through large-scale integration. At this writing, the LSI parts have been available for approximately two years, and are beginning to have a quite noticeable impact on consumer product designs.

An important way to monitor the quality of encoded sound tracks is to look at them on an X-Y oscilloscope. Left, center, and right only information should be seen as straight lines pointing vertical, at 45 degrees, and horizontal respectively (if Left Total is applied to the vertical channel, and Right Total to the horizontal channel). Although it may be difficult to find in all programs, most contain moments where each of the channels becomes dominant so that they can be seen on this test. Pure surround information is shown as a 45 degree line perpendicular to center. A noisy looking sphere results when all output channels are equal in level. Phase errors can be seen at any point in the 2-channel part of the chain by careful watching of center dialog. It should not change angle with level, nor spread into an ellipse, nor become a curved line. Any such change demonstrates differences between the two channels which should be eliminated.

Of course, any system which uses audio amplitude-phase matrix technology to reduce the number of transmission channels cannot produce perfectly discrete results. Also, surround sound decoder design has many issues which must be faced during design as trade-offs between one element of good sound reproduction, and another. Unfortunately, there are at least two problems in evaluating surround sound decoders which makes them difficult to test, and thus for objective criticism to readily sort out the differences among units in the marketplace. These are:

> 1. That instrumentation is complex. Sounds need to be encoded all over the amplitude-phase space with variations in level, transient behavior, and the like fully excercised, and
>
> 2. That even with complex instrumentation and test signal "suites" it is always possible for a piece of pathological program material to exist that causes problems which would not be caught by an instrumented method.

An example for the first is in quoted and measured crosstalk levels. If the crosstalk is measured only at the extreme conditions, called the cardinal ones, say with a left-only signal, very high numbers may be produced which may not be indicative of the general performance. In other words, panning the left-only signal to part-way between left and center may destroy the high separation numbers that are achieved only at the cardinal positions. And this example is only a static one: dynamic considerations greatly complicate decoding issues.

These problems put subjective testing into a prominent role. What sound designers and product designers do is to find program material which shows particular susceptibility to certain problems, and then to run this program material on any newly proposed technology in order to test any newly proposed technology.

Among the inherent problems with matrix encoding is "magic surround." This refers to the fact that any phase discrepancy in the original recording, such as the use of spaced microphone music recordings, will result in a greater level of the corresponding material in the surround than in a discrete system. This has shown up in film exhibition in interesting ways. For example, it has happened that a picture has opened in a secondary market with 35-mm prints, because not enough 70-mm prints were available at first. Following opening, a 70-mm print becomes available and is substituted for the 35-mm one. The theater manager thinks something has gone wrong because the print contains less surround level than the 35-mm one! This is enough to produce a service call, when actually nothing is wrong: it is simply the magic surround problem manifest. During post-production, excessive magic surround is weeded out at an early stage by monitoring through a matrix encoder and decoder for signs of trouble. A slight narrowing of the stereo field of two- or three-channel recordings is usually sufficient to reduce the effect adequately.

A second problem is a mis-positioned bit of sound. There is a tradeoff between fast positional accuracy, and causing audible distortion as control voltage changes become so rapid that the gain is changed over part of a cycle, producing distortion. The tradeoff solution is to make the control voltage speed just somewhat slower than that which causes audible distortion, and this means at about syllabic rate. Let us say that a sound effect occurs in the right channel just before a line of dialog comes in in the center. The dominance control will have essentially "panned" the decoder right, but then the dialog pulls it back to center. A classic case of these occurred in *Star Trek II* where many lines of dialog were preceded by breaths taken by the actors that are heard wholly to one side or the other, whereas the dialog occurs in the center. This is disconcerting to listeners, but better developed professional decoders have reduced the problem in recent years. The way dynamic mis-positioning are avoided is by monitoring mixes through the encoder and decoder so that they can be spotted in early stages and corrected. Correction involves either "ducking" the effect which is doing the pulling just before the sound which is being pulled, or "popping" the level of other sound up in the cross-screen channel.

A third potential problem arises due to the accuracy of the phase shift networks in the encoder. Should they not be at precisely plus and minus 90 degrees from center encoding, then pans between center and surround will pass over either left or right, instead of maintaining the potentially high separation from left and right. This potential problem can be minimized by tight tolerance on the all-pass networks used in the encoder.

Decoders which deviate from standard ones are subject to their own set of problems:

1. Mis-positioning is controlled by monitoring over a standard professional encoder and decoder pair. If the consumer decoder uses different time constants from the dubbing stage one, the corrections put in during mixing may well be wrong and result in audible pulling, particularly of dialog.

2. Another effect which has been heard and may originate with the same source is that sound effects, meant to be on screen, wind up in the surrounds momentarily. This may also be due to differing control algorithms from the one which has been monitored during mixing.

2. The process of control must work well over a wide dynamic range. Low-level errors such as dc-bias errors or drift of circuitry can cause either a lack of steering at low levels, or a bias towards one channel which will make it more pronounced than the others, so low-level sounds will seem to collapse to that channel.

3. Control voltages produced with thresholds or knees can make the steering jump as the threshold is crossed. DSP implementations can suffer from "zipper noise" caused by gain quantization.

4. Saturating time constant capacitors make recovery from loud sounds potentially poor, causing soft sounds following loud ones to be unsteered, and thus vague in localization.

5. Too much center channel preference leads to the apparent width of stereo ambiences and music modulating as dialog comes and goes. This effect can occur if a wideband limiter is used to process Lt and Rt.

6. Any control voltage feedthrough problem in the voltage controlled amplifiers can produce audible thumps.

In addition to these, there is one transformation of the program material from playing in the dubbing stage to playing at home that potentially changes the audibility of crosstalk and decoding errors. In the dubbing stage, relatively large distances are used between the screen and the listeners. This means that the left and right channel loudspeaker signals are delayed slightly with respect to the center channel at listening locations near the center of the console. With typical dubbing stage dimensions and 50 degree wide subtended angles for picture width, this delay is in the range of a few milliseconds. Such a delay is the range in which the precedence effect is fully active, and listeners will tend to perceive the direction as being from the one signal first heard: the crosstalk is effectively suppressed. In the case of the largest theaters, the time delays will become greater, the precedence effect will have less of an impact, and audible crosstalk can become a greater problem, requiring careful decoder design under multiple operating conditions.

In the case of the home, where the distances are all much smaller, a different subjective impression may occur due to crosstalk; since the delay time associated with the crosstalk is now much shorter, the region is encountered wherein timbral change is more prominent than at later delay times. To date, this is more of a theoretical concern rather than a perceived problem.

On the Aesthetic Use of Surround

Despite all the technology described above, the utilization of surround by sound artists is also subject to aesthetic constraints. This occurs because the surrounds are unsupported by picture images, unlike the three screen channels. Also, due to the use of a surround array to get uniform sound pressure over the seating area, certain perceptual constraints occur.

To describe the problem, take the effect of a discrete, clearly audible sound effect in the surround. The audience's attention may be drawn to look at a loudspeaker, which is certainly not desirable—after all, there's a picture on the screen to look at. Further, due to the predominance of the precedence effect, each part of the audience will be drawn to localize to the surround loudspeaker located most closely to them, which obviously varies greatly over the audience. This violates the principle that we are attempting to give the whole audience the same impression of the art on the screen. The best aesthetic use of surround is for two purposes: low-level generally delocalized ambiences used to "pull the audience into the action," and also, the transient moving object headed toward or off the screen.

The perceptual constraints of surrounds involves the lack of timbre matching between the screen channels and the surround channel. Fundamentally this occurs because of the difference in the sound fields generated by the generally directional single-source screen channel over the generally non-directional, and multiple loudspeaker surround channel. In a series of experiments, I found that screen channel and surround channel timbre matching on difficult signals is impossible in theatrical installations due to the difference between the sources. Comb filtering due to the multiple surround sources predominates timbre so much that 1/3-octave frequency response is inadequate to describe the differences between screen and surround channels.

Introduction into the Home

Home video media generally uses pairs of tracks, often duplicated in several different technologies. That is, although VHS for example is equipped with both two longitudinal audio tracks recorded conventionally and with "Hi-Fi" tracks recorded along with the video, the two pairs contain virtually identical information, with the Hi-Fi tracks being simply a better quality reproduction scheme. Thus, only two tracks are available to home video media decoders, usually drawn from the highest available quality ones.

By far the most prominent method of recording is to transfer the theatrical mix released in the conventional Dolby Stereo format "straight across," that is, without any additional signal processing. The Dolby Stereo Print Master (with A-type NR, that is, not the SR Print Master should one exist—the A-type NR characteristic is decoded during transfer) has already had its volume range limited for the limited dynamic range of the optical sound track, and thus is generally suitable for the Hi-Fi tracks of VHS, SVHS, and the digital audio channel of the Laserdisc. (Longitudinal tracks may require some additional limiting, but we are mainly concerned with the best quality tracks here.) In fact, in the best cases, with the analog and digital tracks available on Laser Disc, the mass production process can be compared in an instantaneously switched, blind comparison, with its source master, and the original personnel of the film listening on a dubbing stage aligned to all of the standards, can find the mass production copies audibly perfect, that is, indistinguishable from the master film.

So in the best of all possible cases, it is possible to recover the audio from the medium, and to decode it into four channels, in a way which is audibly identical to the process occurring on the dubbing stage during the making of the film. Video releases with matrix encoding are labled with Dolby Stereo or Ultra Stereo logos so that the user knows to engage the decoder.

The Fourth Channel: Center?

Among the four channels, left, center, right, and surround, it may seem that home systems could most easily give up the center, since stereo home systems work apparently well with phantom center images. While this may be the general perception of stereo, in fact, the "hard" center channel has some obvious, and some not so obvious advantages over the phantom center system:

> 1. If a center channel signal is applied to a center channel loudspeaker, and the crosstalk is low, then everyone, despite where they sit, will perceive the direction of the center sounds as being from the center of the screen. This enormously opens up the stereo "sweet spot" with regards to center images, and reduces the error in the rest of the stereo sound field dramatically.
>
> 2. A-B comparisons of well aligned two-channel vs. three-channel systems at equal loudness and with matched timbre reveals that the 3 channel system is preferable on "clarity," even when sitting on the precise centerline and not moving one's head. There is simply no better word to describe the result of the A-B comparison: it is as though understanding speech in the presence of interfering noise from other elements of the sound track is simply easier, although "intelligibility" remains the same. The presumption is that the phantom image system requires greater attention to perceive the phantom,

and that 3-channel listening to center is more like natural stereo listening—enough more so that it is "easier" to listen to.

Conclusion

Surround sound technology has been explored in use. Some of the considerations which mixers need to have in order to best use the technology have been discussed. It should be recognized that although this paper discusses potential drawbacks in many areas, surround sound technology is the only practical way to experience an important dimension, the spatial one, for many release formats, and that there are many examples of competent sound artists using the medium to great artistic effect, with the result being commercial success in the marketplace.

---

1. Johnston, William A., "The Public and Sound Pictures," Transactions of the Society of Motion Picture Engineers, vol. XII, no. 35, meeting of September 24-28, 1928, p. 614. Mr. Johnston was the editor of Motion Picture News.

2. ibid.

3. Motion Picture Sound Engineering, Academy of Motion Picture Arts and Sciences, 1938.

4. ANSI Standard SMPTE A202M